

Keuren is mensenwerk : kwaliteitsaspecten van de medische keuring voor militairen binnen de Nederlandse krijgsmacht en voor dienstplichtigen in het bijzonder

Citation for published version (APA):

Reulings, P. G. J. (1996). *Keuren is mensenwerk : kwaliteitsaspecten van de medische keuring voor militairen binnen de Nederlandse krijgsmacht en voor dienstplichtigen in het bijzonder*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. <https://doi.org/10.26481/dis.19961212pr>

Document status and date:

Published: 01/01/1996

DOI:

[10.26481/dis.19961212pr](https://doi.org/10.26481/dis.19961212pr)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 05 May. 2023

SAMENVATTING

Doelstelling van dit proefschrift is het zoeken naar verbeteringen van de kwaliteit van de medische keuringen binnen de Nederlandse krijgsmacht. Bijzondere aandacht is geschonken aan dienstplichtkeuringen en keuringen voor vrijwillig dienende militairen van de Koninklijke Landmacht.

Ofschoon er door de opschorting van de opkomstplicht per 1 januari 1997 geen dienstplichtkeuringen meer worden verricht, is dit onderzoek toch relevant. Dienstplichtkeuringen kunnen in de toekomst weer actueel worden als de Dienstplichtwet door calamiteiten of oorlogsdreiging in werking treedt. Daarnaast vormen de eisen en procedures van dienstplichtkeuringen de basis van alle militairen keuringen (o.a. aanstellingskeuringen en periodiek geneeskundige onderzoeken binnen de krijgsmacht). De resultaten zijn daardoor van belang voor al deze medische keuringen binnen de krijgsmacht. Suggesties ter verbetering van het keuringsapparaat zijn derhalve krijgsmachtbreed.

De dienstplichtkeuring is te beschouwen als een bijzondere vorm van de aanstellingskeuring. In de burgermaatschappij staan aanstellingskeuringen in de volle belangstelling. Het gaat daarbij vooral over de effectiviteit van de aanstellingskeuring als instrument voor selectie en risicowering. Aanstellingskeuringen zijn alleen wenselijk of noodzakelijk voor functies waar de persoon in kwestie of derden belangrijke gezondheids- en/of veiligheidsrisico's loopt/lopen. En dit geldt ook voor militaire functies.

In **hoofdstuk 1** wordt de betekenis van (aanstellings)keuringen en in het bijzonder van keuringen binnen de krijgsmacht beschreven (**paragraaf 1.1**).

Uniek aan militaire keuringen is dat, in tegenstelling tot vele aanstellingskeuringen in de burgermaatschappij, gedetailleerde richtlijnen gebruikt worden. De keuringsuitslag wordt beoordeeld aan de hand van het militair keuringsreglement (MKR) en het Aanwijzingenboek voor artsen die belast zijn met militaire keuringen. De arts dient niet alleen de keuringsuitslag vast te stellen; ook dienen aanwezige ziekten en gebreken volgens de codering van de International Classification of Disease (ICD) te worden vastgelegd. Bij een geschiktheidverklaring kent de arts een ABOHZIS-classificatie aan de persoon toe. De ABOHZIS-classificatie dient als hulpmiddel bij het indelen naar functie van de goedgekeurden. Elke letter van de ABOHZIS staat voor een bepaalde rubriek: A= Algemeen fysieke toestand; B= Bovenste ledematen; O= Onderste ledematen en rug; H= Horen; Z= Zien; I= Intelligentie en S= (Psychische) Stabiliteit. Bij het constateren van ongeschiktheid dient de arts de reden van afkeuring aan te geven in een code conform een lijst van het MKR en de rubriek van de ABOHZIS-classificatie waarop de afkeuring betrekking heeft.

De actuele mogelijkheden om de kwaliteit van de militaire keuringen te onderzoeken worden vervolgens in **paragraaf 1.2** besproken. De belangrijkste aspecten om de kwaliteit van de keuring te meten zijn reproduceerbaarheid en validiteit.

De reproduceerbaarheid geeft aan in hoeverre de keuring dezelfde uitslag oplevert, indien deze op meerdere momenten plaatvindt of door andere artsen wordt verricht.

Bij de militaire keuring kan deze worden bepaald door na te gaan in welke mate de keurend artsen overeenstemmen in de beoordeling van de medische geschiktheid voor militaire dienst.

Reproduceerbaarheid is een voorwaarde van een valide keuring. De validiteit geeft aan in hoeverre met de keuring de geschiktheid van de gekeurden juist beoordeeld wordt. Door het ontbreken van de gegevens over de afgekeurden, zijn de mogelijkheden om de validiteit van de keuring te bepalen beperkt. Alleen de mate waarin dienstplichtigen in dienst om medische redenen worden afgekeurd kan worden nagegaan. Dit geeft een indicatie van de voorspellende waarde van de keuring. Prognostische factoren van dit medisch disfunctioneren kunnen worden nagegaan, zodat aan de hand van risicoprofielen het medisch disfunctioneren beter kan worden voorspeld.

Het onderzoek is toegespitst op de voorspellende waarde en de reproduceerbaarheid van de keuring. Dit heeft geleid tot de volgende vragenstellingen:

- In hoeverre is het mogelijk het medisch disfunctioneren in militaire dienst te voorspellen?
- In welke mate zijn militaire keuringen reproduceerbaar?

In **hoofdstuk 2** wordt de eerste vraagstelling beantwoord met behulp van dienstplichtgegevens. In **paragraaf 2.1** wordt het keuringstraject van de dienstplicht beschreven. Een groep van 118.716 dienstplichtigen wordt van het begin (vanaf 1987) tot het einde van dit traject gevolgd. Ondanks de uitgebreide medische selectie disfunctioneert zes procent van deze dienstplichtigen in diensttijd bij de Koninklijke Landmacht om medische redenen.

Nader onderzoek naar de medische disfunctioneerders is beschreven in **paragraaf 2.2**. De resultaten van het retrospectieve cohortonderzoek bij 8714 dienstplichtigen die in 1989 de eerste dienstplichtkeuring hebben ondergaan en in 1990 bij de Koninklijke Landmacht zijn opgekomen, worden weergegeven. De meeste disfunctioneerders blijken ongeschikt te zijn verklaard wegens aandoeningen van de onderste extremiteiten inclusief de rug of onvoldoende psychische stabiliteit.

Met behulp van logistische regressieanalyse is gezocht naar risicoprofielen die deze twee soorten van disfunctioneren kunnen voorspellen voordat de dienstplichtigen in dienst komen. Prognostische factoren van het disfunctioneren wegens aandoeningen van de onderste extremiteiten en de rug zijn: het constateren van een functiebeperking tijdens de eerste keuring op grond van aandoeningen van de onderste extremiteiten en de rug, die in eerste instantie geen afkeuring leek te rechtvaardigen of het ondergaan van een herkeuring omdat men bij de eerste keuring tijdelijk ongeschikt is verklaard en/of het ondergaan van een extra keuring naar aanleiding van de gegevens ingevuld op het anamneseformulier vlak vóór opkomst.

Prognostische factoren van het disfunctioneren vanwege onvoldoende psychische stabiliteit zijn: het hebben van een psychische stoornis bij de eerste keuring die niet tot een afkeuring leek te hoeven leiden, een bipatride nationaliteit en terecht komen in die lichtingsploegen, waar het aanbod van dienstplichtigen kleiner is dan de overige lichtingsploegen. Met behulp van de, op de prognostische factoren gebaseerde, risicoprofielen kan de kans om te disfunctioneren wegens aandoeningen van de onderste extremiteiten inclusief de rug of onvoldoende psychische stabiliteit vóór opkomst worden berekend. Als bij de keuring dit extra selectiemiddel gehanteerd zou worden en dienstplichtigen met een verhoogd risico niet in dienst komen, kan het aantal medische disfunctioneerders in beperkte mate worden verminderd. De uitkomsten van dit onderzoek hebben geleid tot een discussie in het Tijdschrift Sociale Gezondheidszorg over de voor- en nadelen van deze selectiemethode. In **bijlage 1, 2 en 3** is deze discussie letterlijk weergegeven.

Hoofdstuk 3 geeft de resultaten van onderzoeken die verricht zijn om de reproduceerbaarheid van militaire keuringen te achterhalen.

Allereerst worden in **paragraaf 3.1** de gegevens van de dienstplichtkeuringen van 1993 geanalyseerd. Iets meer dan de helft van de afkeuringen blijkt betrekking te hebben op aandoeningen die vallen onder de algemene fysieke toestand of aandoeningen van de onderste extremiteiten inclusief de rug. Er blijkt aanzienlijke variatie te zijn wat betreft de beoordeling van de medische geschiktheid van dienstplichtigen. Dit geldt zowel tussen de indelingsraden als tussen de artsen. De grootste variatie is aanwezig bij de afkeuringen op grond van aandoeningen van de onderste extremiteiten en de rug. De variatie is deels te verklaren door procedureverschillen tussen de indelingsraden met betrekking tot het afkeuren wegens onvoldoende gehoor- en gezichtsvermogen. De overige inter-raadvariatie is het gevolg van verschillen tussen de individuele artsen.

In **paragraaf 3.2** wordt een experiment beschreven naar de mate van overeenstemming tussen artsen omtrent de medische geschiktheid van dienstplichtigen voor militaire dienst. Een 36-tal artsen hebben bij 14 geselecteerde 'moeilijk beoordeelbare' casussen (met betrekking tot aandoeningen van de onderste extremiteiten en de rug) de geschiktheid van dienstplichtigen beoordeeld. De mate van overeenstemming tussen artsen is bepaald met behulp van Cohen's kappa. De kappawaarde kan variëren tussen 0 en 1, waarbij een waarde van 0 tot 0.40 impliceert dat er een slechte overeenstemming is. Zowel de overeenstemming bij de keuringsuitslag als bij de beoordeling van de beperkingen ten aanzien van onderste extremiteiten en de rug is slecht: de kappa bedraagt respectievelijk 0.20 en 0.17. De overeenstemming tussen de artsen binnen één indelingsraad is ook laag: tussen de indelingsraden schommelt de kappa tussen 0.19 en 0.40 betreffende de keuringsuitslag en tussen 0.13 en 0.35 betreffende beperkingen door aandoeningen van de onderste extremiteiten en de rug. Intercollegiaal overleg en de sturing door het hoofd van de artsen op de indelingsraad hebben niet een zodanig overwicht dat de artsen van eenzelfde raad hetzelfde beleid voeren. Tevens lijkt contextuele informatie de overeenstemming tussen artsen te verlagen: bij casussen met deze informatie bedraagt de kappa ten aanzien van de keuringsuitslag 0.07, terwijl die bij casussen zonder contextuele informatie 0.35 is. Wat betreft de beperkingen door aandoeningen van de onderste extremiteiten en de rug is dit 0.07 ten opzichte van 0.21.

Ter verbetering van de overeenstemming zou men kunnen overwegen functiegericht te keuren. Een aanvullende mogelijkheid is dat artsen zich specialiseren in het beoordelen van *specifieke* functies of functiegroepen. Daarnaast zou meer aandacht besteed kunnen worden aan intercollegiaal overleg. Trainingen zijn zinvol voor keurend artsen. Dat geldt vooral voor degenen die de artsen inwerken.

In **paragraaf 3.3** wordt een aanvullend onderzoek beschreven naar de overeenkomst tussen artsen die periodiek geneeskundig onderzoek (PGO's) verrichten bij beroepspersoneel van de Koninklijke Landmacht. Bij dit onderzoek zijn 24 aselect gekozen casussen voorgelegd aan 2 artsen. De overeenkomst tussen beoordelingen van de oorspronkelijke arts en de beoordelingen van de twee andere artsen is vastgesteld aan de hand van Cohen's kappa. Bij deze keuringen geven de artsen niet alleen de keuringsuitslag en de ABOHZIS-classificatie, maar ook de medische beperkingen aan. De overeenkomst ten aanzien van de keuringsuitslag is onder te verdelen naar de beoordeling van de A-rubriek (de algemene fysieke toestand), de

B-rubriek (aandoeningen van de bovenste extremiteiten) en de O-rubriek (aandoeningen van de onderste extremiteiten en de rug). De kappa bedraagt bij de beoordeling van de A-, B- en O-rubriek respectievelijk 0.16, 0.52 en 0.44. Bij het wel of niet toekennen van een medische beperking is de kappa achtereenvolgens voor de A-, B- en O-rubriek 0.67, 0.74 en 0.65. Ook in dit onderzoek blijken de keuringsuitslagen op grond van de ABOHZIS-classificatie tussen individueel keurend artsen onvoldoende reproduceerbaar. De keuringsuitslag is derhalve niet alleen afhankelijk van de status praesens van de militair, maar ook van de keurend arts. Het gebruik van de ABOHZIS-classificatie levert geen meerwaarde op bij het periodieke onderzoek naar de dienstgeschiktheid van de individuele militair.

In de beschouwing (**hoofdstuk 4**) is nagegaan in hoeverre de vraagstellingen beantwoord zijn. Onderzoek heeft uitgewezen dat het medisch disfunctioneren op grond van de O- en S-rubriek deels kan worden voorspeld aan de hand van risicoprofielen. Het aantal medische disfunctioneerders zal door deze extra selectie kunnen verminderen. Echter een beperkt deel van de medische disfunctioneerders is onderzocht en de risicoprofielen hebben niet een zodanige impact dat de voorspellende waarde (van 91,9% naar 92,3%) aanzienlijk verhoogd wordt.

De overeenkomst tussen de artsen bij de beoordeling van schriftelijke dossiers is slecht. Dit geldt ten aanzien van de beoordeling van de keuringsuitslag zowel bij dienstplichtkeuringen als bij PGO's van vrijwillig dienende militairen van de Koninklijke Landmacht. Voldoende reproduceerbaarheid van de keuringsuitslag van de militaire keuring kon niet worden vastgesteld, ondanks de richtlijnen die in de MKR en het Aanwijzingenboek zijn weergegeven om de keuring zo uniform mogelijk te laten verlopen. De arts beschikt, bij de beoordeling van de geschiktheid voor de militaire dienst, over een bepaalde mate van beslissingsruimte. Daardoor blijft keuren mensenwerk.

Naar aanleiding van de resultaten van het onderzoek zijn suggesties gedaan ter verbetering van de militaire keuringen, zoals de dienstplichtkeuringen, aanstellingskeuringen en periodieke keuringen voor beroepsmilitairen, die allen volgens hetzelfde basisprincipe werken. De belangrijkste aanbeveling is dat ook bij militaire keuringen functiegericht dient te worden gekeurd. Bij deze optie zou men ook kunnen overwegen om artsen zich te laten specialiseren in de beoordeling van specifieke functiegroepen, zoals chauffeurs of infanteristen. Door het functiegericht keuren heeft het ABOHZIS-classificatiesysteem, mede door de te geringe reproduceerbaarheid, geen waarde meer voor de krijgsmacht 'nieuwe stijl'. Binnen het principe van het functiegericht keuren past de ontwikkeling van het functiegram-somagram goed, omdat voor elke persoon de belasting van de functie wordt afgezet tegen de belastbaarheid van de persoon in kwestie.

Om in te spelen op de vernieuwingen binnen de krijgsmacht dient verder onderzoek te worden verricht naar de keuringen van het 'nieuwe' beroepspersoneel. Allereerst kan de mate van medisch disfunctioneren worden nagegaan. Zeker als het keuringssysteem verandert of de functie-eisen verzaamd of verlicht worden is het aan te bevelen het medisch disfunctioneren systematisch te evalueren. Deze evaluatie vormt een uitstekende graadmeter voor de voorspellende waarde van de keuring. Daarnaast kan worden onderzocht of het systeem van functiegericht keuren ten goede komt aan de reproduceerbaarheid.

In navolging van eerder onderzoek kan het effect hiervan beter met simulatiekeurlingen gedaan worden dan met schriftelijke dossiers.

Tijdens het onderzoek is steeds weer opgevallen dat over de gehele Nederlandse mannelijke bevolking over een lange periode keuringsgegevens aanwezig zijn. De vraag is gerezen of deze gegevens ook voor een ander doel gebruikt kunnen worden dan waarvoor zij primair zijn bedoeld, namelijk als basis voor onderzoek naar de gezondheidstoestand van de 18-jarige Nederlandse mannen. Gebleken is dat in het verleden hieromtrent onderzoek is verricht, mede door de mogelijkheid om de gegevens te koppelen aan externe bestanden. Ofschoon nu geen dienstplichtkeuringen meer worden verricht en de privacyregelingen koppeling met externe bestanden niet toelaat, is het altijd mogelijk om retrospectief onderzoek te doen naar trends van belangrijke aandoeningen. Het zou betreurenswaardig zijn als deze waardevolle bron van informatie, door het verdwijnen van de dienstplichtigen, niet gebruikt wordt.

SUMMARY

The purpose of this thesis is to determine the quality of the medical examination within the Dutch armed forces. Special attention is given to the examination of conscript and voluntary personnel of the Royal Netherlands Army (RNLA).

Despite the fact that conscription is being suspended in the Netherlands from 1 January 1997, this investigation is still relevant. The examination of conscripts could be reintroduced in times of a national calamity, or threat of war. Furthermore, the requirements and procedures of conscript examination form the basis of all military medical examinations.

Therefore, the results of the investigation and suggestions for improvement of the medical examination apparatus are important for all medical examinations within the armed forces. Conscripts' medical examination can be considered as a special form of pre-employment medical. In the Netherlands the issue of pre-employment medicals enjoys wide public interest. This discussion addresses the effectiveness of the pre-employment medical as an instrument for selection and avoiding risks. Pre-employment medicals are only recommended or necessary for jobs in which, as is the case in military functions, the health or safety risks are large.

Chapter 1 describes the meaning of the term (entrance) medical examination with particular reference to the military context (**paragraph 1.1**).

In comparison with civilian medical examinations, military medical examinations are unique in that they are conducted on the basis of detailed written guidelines.

Examination results are judged according to the Regulations for Military Medical Examinations (abbreviated to MKR in Dutch) and are described in terms of the Book of Medical Indications available for physicians charged with this task.

Doctors not only determine a person's fitness; illnesses and disorders are recorded according to the International Classification of Diseases (ICD codes).

For the purposes of a declaration of fitness physicians classify persons undergoing a medical examination according to the ABOHZIS classification system. The ABOHZIS classification serves as an aid for assigning military functions.

Each of the letters making up ABOHZIS stands for a certain category: A: general physical condition; B: upper extremities; O: lower extremities and back; H: hearing; Z: sight; I: intelligence; S: mental stability.

If a candidate is declared unfit for military service, physicians are required to state the reason by indicating a code in accordance with the MKR list and the ABOHZIS classification.

The technical scope for assessing the quality of military medical examination is covered in **paragraph 1.2**. The most important aspects to measure the quality of the medical examination are reproducibility and validity. The reproducibility indicates to what extent the examination produces the same results if held at different moments, or performed by different physicians. In the case of military medical examinations this aspect may be determined by establishing to what extent the physicians performing the examinations agree in their assessment of conscripts' medical fitness for military service.

Reproducibility is a precondition for a valid medical examination. Its validity indicates to what extent the medical examination can be used to assess correctly the fitness of those undergoing

the examination. Through lack of information concerning those who fail, the possibilities for assessing the validity of an examination are limited. It is only possible to establish the extent to which conscripts on duty are discharged for medical reasons. This gives an indication of the predictive value of the medical examinations. Predictors of this medical dysfunctioning can be established, which enables predictions of medical dysfunctioning to be improved, using risk profiles. The investigation focuses on the predictive value and the medical examination's reproducibility, which led to the following questions:

- To what extent is it possible to predict medical dysfunctioning in military service?
- To what extent are military medical examinations reproducible?

On the basis of data from the medical information system, the first question is answered in **chapter 2. Paragraph 2.1** describes the possibilities for medical examination and re-examination of conscripts.

A group of 118,716 conscripts was followed from the beginning (1987) to the end of this course of (re-)examination. Despite extensive medical selection, 6 percent of these conscripts on duty in the RNLA are discharged for medical reasons.

Further investigation into medical dysfunctioning is described in **paragraph 2.2**. It contains the results of the retrospective cohort investigation of 8,714 conscripts who underwent a medical examination in 1989 and a year later in 1990 when they were called up for military service in the RNLA.

Most of those who were discharged from military service were declared unfit because of disorders of the lower extremities and the back, or because of insufficient mental stability. Using logistical regressive analysis, risk profiles were sought that can predict these two sorts of dysfunctioning before conscripts enter military service.

Predictors of dysfunctioning due to disorders of the lower extremities and back include: establishing a functional restriction during the first medical examination on the grounds of disorders of the lower extremities and the back which did not initially seem to justify rejection on medical grounds; or undergoing a medical re-examination, after having been found temporarily unfit during the first medical examination; and/or undergoing an extra examination on the basis of a case history form immediately prior to conscripts' actual enlistment. Prognostic factors of dysfunctioning due to a lack of mental stability include establishing a mental disorder during the first medical examination which did not seem to have to lead to rejection on medical grounds; and/or having dual nationality; and/or ending up in drafts with a smaller supply of conscripts than in other drafts.

Profiles, based on these predictors, can be used to calculate the risk of conscripts' dysfunctioning owing to disorders of the lower extremities (including the back) or insufficient mental stability before they are called up for military service.

It appears that if such an extra selection method is applied, the number of the conscripts who are discharged for medical reasons can to some extent be reduced.

In response to this investigation, there has been a discussion on the advantages and disadvantages of this selection method.

This discussion was held in the 'Tijdschrift Sociale Gezondheidszorg' (Journal of Social Health) and has been incorporated in **Appendices 1, 2 and 3**.

Chapter 3 shows the results of research which was done to investigate the reproducibility of medical examination in the RNLA.

Firstly, in **paragraph 3.1** the data of the 1993 medical examinations are analyzed. Just over half of the conscripts who fail the medical examination do so because of disorders that fall under the categories of general physical condition or disorders of the lower extremities (including the back).

Variation can be observed between the different examination boards as well as between the physicians who conduct these examinations. This variation is reflected most clearly in the assessment of conscripts rejected due to disorders of the lower extremities and the back.

The variation can be partly explained due to procedural differences between examination boards with regard to candidates' varying sight and hearing abilities.

The remaining variation between the different boards is due to the differences between individual medical examiners.

Paragraph 3.2 describes an investigation into the degree of consensus among physicians concerning the medical fitness of conscripts.

Some 36 physicians reviewed the records of 14 cases in which it was difficult to assess the fitness because of disorders of the lower extremities and the back.

Agreement was calculated by means of Cohen's kappa. The kappa value can range from 0 to 1. A value of 0-0.40 indicates poor agreement. The agreement on fitness, as well as on the classification with regard to the lower extremities and the back are low, with kappa values of 0.20 and 0.17 respectively.

The influence of the conscription board on inter-observer agreement appears to be small, because the agreement between physicians within a single conscription board varies between 0.19 and 0.40 where candidates' fitness is concerned and between 0.13 and 0.35 for the classification of the lower extremities and the back.

Consultation between colleagues and steering by the head of the medical commission does not result in physicians of the same board consistently applying the same policy. It also appears that contextual information has a negative effect on inter-observer agreement. In the assessment of candidates' fitness, kappa is 0.07 for cases with contextual information, whereas this value was 0.35 for cases without contextual information. When we look at the classification of the lower extremities and the back, the kappa values were 0.07 and 0.21 respectively. To improve inter-observer agreement the idea of carrying out job-specific medical examinations might be considered. Another idea would be for the physicians to specialize in the examination of candidates for *specific* functions or groups of functions. Furthermore, more attention could be given to consultation among colleagues and to courses for medical examiners, especially those 'breaking in' new physicians.

An additional investigation is described in **paragraph 3.3**, which relates to the degree of agreement among physicians who carry out periodical medical examinations of regular RNLA personnel. In this study 12 randomly selected cases were presented to two physicians. The agreement between the original physician and the other two is established on the basis of Cohen's kappa. In these medical examinations the physicians not only gave an assessment of candidates' fitness and the ABOHZIS classification, but they also indicated medical restrictions. With respect to the ABOHZIS classification it can be observed that in the A category (general

physical condition) kappa was 0.16, in the B category (upper extremities) 0.52 and in the O category (lower extremities and back) 0.44. The kappa values for the medical restrictions indicated in the A, B and O categories were 0.67, 0.74 and 0.65 respectively.

This investigation also shows that the inter-observer agreement on the assessment of the fitness between the physicians is difficult to reproduce. The assessment of people's fitness therefore not only depends on their general health condition, but also on the examiner. The use of the ABOHZIS system does not contribute to periodical medical examinations of individual service personnel.

In the discussion (**chapter 4**) an attempt was made to establish to what extent the questions that were presented have been answered.

This investigation has shown that conscripts' medical dysfunctioning based on the O and S categories can be predicted, using risk profiles. The number of conscripts who dysfunction for medical reasons can therefore be reduced by this extra selection. However, only part of those who dysfunction for medical reasons were examined and the risk profiles did not lead to drastic improvements in the predictive value (from 91.9% to 92.3%).

On the basis of written files it can be established that there was poor to moderate agreement between the physicians. This conclusion can be drawn after reviewing examination results of both conscripts and RNLA regular personnel. Despite the regulations and written guidelines the reproducibility of the military medical examinations appears to be limited. The scope for physicians to decide for themselves is an important feature in assessing someone's medical fitness for military service, for which the human factor is thus indispensable.

As a result of this investigation suggestions were made to improve the quality of the military medical examinations, including medical examinations for conscripts and regular personnel, as well as pre-employment medical examinations, all of which share the same basic principle. An important recommendation is made, namely to develop a system in which the military medical examinations are related to specific job descriptions. In such a system physicians could specialize in the examination for certain functions, such as infantrymen or drivers.

In a job-related system to be further developed in the 'new' all-volunteer armed forces there will be no need for the ABOHZIS system, with its limited reproducibility.

The development of the functiongraph-somagraph fits well within the principle of job-related examination, because the physical demands of someone's job are compared to his/her physical condition.

Because of the reorganisation of the armed forces, further investigation is needed of the medical examination system of the 'new' regular personnel. Firstly, research in medical dysfunctioning within the professional army is necessary. If the examination system changes or if functional demands increase or decrease, dysfunctioning for medical reasons needs to be evaluated on a systematic basis, because this can serve as an indicator of the predictive value of medical examinations. Furthermore, research has to be done to investigate if job-related examinations yield higher reproducibility. The effect can be measured with simulating persons instead of written case histories.

Throughout the investigation it was repeatedly noted that examination data exist with respect to the entire Dutch male population. The question arose as to whether these data can also be used for other than their primary purposes, namely as a basis for an investigation into the state of health of Dutch men of the age of 18. It has appeared that studies were made in this areas in the past, partly as a result of the possibilities to combine these data with external databases. Although no medical examinations of conscripts are currently being carried out and privacy regulations do not allow external databases to be linked, retrospective research can always be carried out into the trends of the main disorders. It would be regrettable if, owing to the disappearance of conscription, this valuable source of information not to be used anymore.